The Canadian Journal of Statistics
Vol. 34, No. 4, 2006, Pages ???-???
La revue canadienne de statistique

Nonparametric Test for Checking Lack-of-Fit of Quantile Regression Model under Random Censoring

Lan WANG

Key words and phrases: conditional moment test; lack-of-fit test; model checking; random censoring; quantile regression.
MSC 2000: Primary 62G10; secondary 62G20, 62N01.

Abstract: Recently, considerable attention has been devoted to quantile regression under random censoring in both statistical and econometrical literature yet little has been done on the important problem of model checking. This paper proposes a nonparametric test for checking the lack-of-fit of the quantile function of the survival time given the covariates when the survival time is subjected to random right censoring. The test statistic is a kernel-based smoothing estimator of a moment condition. The test has an asymptotic normal distribution under the null hypothesis. Its power property is investigated through its behavior under local alternative sequence. Simulation results demonstrate the satisfactory performance of the proposed test in finite sample size. The new test is also illustrated by an application to the Stanford Heart Transplant data.

Title in French: we can supply this

Résumé : Recently, considerable attention has been devoted to quantile regression under random censoring in both statistical and econometrical literature yet little has been done on the important problem of model checking. This paper proposes a nonparametric test for checking the lack-of-fit of the quantile function of the survival time given the covariates when the survival time is subjected to random right censoring. The test statistic is a kernel-based smoothing estimator of a moment condition. The test has an asymptotic normal distribution under the null hypothesis. Its power property is investigated through its behavior under local alternative sequence. Simulation results demonstrate the satisfactory performance of the proposed test in finite sample size. The new test is also illustrated by an application to the Stanford Heart Transplant data.

1. INTRODUCTION

Censored quantile regression provides an appealing semiparametric approach to directly model the effects of a set of covariates on the survival time. Let T_i be the survival time (or a transformation thereof) that depends on an m-dimensional vector of covariates X_i, and let C_i denote the censoring time, $i = 1, \ldots, n$. The survival time T_i is not observed or censored if it is greater than the corresponding censoring time. More specifically, the censored quantile regression model assumes

$$
\begin{aligned}
T_i &= g(X_i, \beta) + \epsilon_i, &(1)\\
Y_i &= min(T_i, C_i), &(2)
\end{aligned}
$$

where $g(\cdot)$ is the conditional quantile function of T_i given X_i, which is known up to a finite-dimensional parameter β, and the ϵ_i's are independent random errors which satisfy $P(\epsilon_i \leq 0|X_i) = \tau$, for some specified $0 < \tau < 1$. For the censored median regression, also called censored least absolute deviation regression, τ takes the value 0.5. Let $\Delta_1, \ldots, \Delta_n$ represent the censoring indicators, i.e., $\Delta_i = I(T_i \leq C_i)$, where $I(\cdot)$ is the indicator function. The observed data thus consist of the triples (Y_i, X_i, Δ_i), $i = 1, \ldots, n$.

Compared to the traditional Cox proportional hazards regression model for survival analysis, the censored quantile regression directly models the survival time, thus is easier to interpret. By varying the value of τ, quantile regression produces a more complete picture of the relationship between the survival time and the covariates. Moreover, the quantile regression can incorporate heterogeneity that can not be accommodated by the Cox model. As Portnoy (Tableman and Kim, 2000, section 8.4) suggested, "The use of both approaches (meaning the censored quantile regression model and the Cox model) provides not only a broader perspective, but is much less likely to miss important relationships."

Ying, Jung and Wei (1995) appeared to be the first to study median regression under random censoring. They observed that the median is a simple and meaningful measure for the center of a long-tailed survival distribution; and that unlike the mean, the median can be well estimated as long as the censoring is not too heavy. Recently, increasing attention from both statisticians and econometricians has been paid to quantile regression under random censoring, see Yang (1999), Honoré, Khan and Powell (2002), Chen and Khan (2001), Subramanian (2002), Portnoy (2003), Qin and Tsao (2003), Gannoun et al. (2005), Huang, Ma and Xie (2005), among others. Tableman and Kim (2000, Chapter 8, contributed by Portnoy) offers a useful tutorial of related methodologies and software application.

In practice, the conditional quantile function is often taken to be linear for its computational convenience and easy interpretation. It is widely recognized that a misspecified quantile function might lead to biased estimator and unreliable inference. The above mentioned work all focuses on the issue of estimation when the conditional quantile function in (1) assumes a known parametric form. Little has been done on the important problem of model checking. Ying, Jung and Wei (1995) recommended a useful graphical technique for model checking based on a plot of a process constructed from cumulative sums of median residuals. The graphical method provides a helpful visual aid but may yield inconclusive indication of lack-of-fit.

There exists a large amount of literature on lack-of-fit tests for mean regression model. These procedures are not directly extendable to quantile regression without serious technical work. The technical challenge is mainly due to the fact that the residuals from quantile regression involve indicator functions that are not differentiable in the regression parameter. Techniques such as Taylor expansion, which are essential for establishing the asymptotic theory of the lack-of-fit tests for mean regression, are not applicable. There have been some papers on lack-of-fit tests for quantile regression without censoring, such as Zheng (1998), Horowitz and Spokony, V. G. (2002), and He and Zhu (2003), but their results do not cover the censored case. Wang (2005) proposed an ANOVA (analysis of variance) type test for censored median regression model when all the censoring variables are observable. This excludes random censoring, where the censoring variables are observed only when censoring happens.

In this paper, we propose a kernel-smoothing based nonparametric test for checking the lack-of-fit of a postulated conditional quantile function versus a general smooth alternative when the survival time is under right random censoring. This test generalizes the work of Zheng (1998) to censored data. Due to the need to estimate the censoring distribution, the derivation of the asymptotic distribution is significantly more complex and involves handling third-order U-statistics. We introduce the test statistic in Section 2. Its asymptotic properties are presented in Section 3. In Section 4, we report simulation results to demonstrate the finite sample performance of the proposed test. In Section 5, we apply the test to the Stanford Heart Transplant data. We conclude with some brief discussions in Section 6.

2. TEST STATISTIC

We are interested in testing whether the τth conditional quantile function of the survival time T_i given X_i can be adequately modeled by the assumed parametric functional form in (1). The null hypothesis is

$$H_0 : P(T_i \le g(X_i, \beta_0)|X_i) = \tau, \quad \text{for some } \beta_0 \in B \tag{3}$$

where B is the parameter space, and the alternative hypothesis is

$$H_1 : P(T_i \le g(X_i, \beta)|X_i) \ne \tau, \quad \text{for all } \beta \in B. \tag{4}$$

The test statistic we introduce below is a nonparametric estimator of a population moment condition which is zero under the null hypothesis and strictly positive for any alternative. To describe the test statistic, assume that the censoring variable C_i has cumulative distribution function $1 - G(\cdot)$ and that the C_i and T_i are independent. We first note that under the assumptions given in Section 3,

$$\begin{aligned} & E(I(Y_i - g(X_i, \beta_0) \ge 0)|X_i) \\ = \;& P(T_i \ge g(X_i, \beta_0)|X_i)P(C_i \ge g(X_i, \beta_0)|X_i) \\ = \;& (1 - \tau)G(g(X_i, \beta_0)). \end{aligned}$$

Define $e_i = I(Y_i \ge g(X_i, \beta_0)) - (1 - \tau)G(g(X_i, \beta_0))$, then $e_1, \ldots, e_n$ are independent with mean zero conditional on the covariates. The e_i's are the building blocks of our test statistic. Let $p(\cdot)$ be the joint probability density function of the m-dimensional covariate X_i, then

$$E[e_i E(e_i|X_i)p(X_i)] = E\{[E(e_i|X_i)]^2 p(X_i)\} \ge 0,$$

where in the left-hand side the inner-layer expectation is with respect to the conditional distribution of $e_i|X_i$, and the outer-layer expectation is with respect to the joint distribution of (X_i, e_i). Note that the equality holds only when H_0 is true. Based on this observation, we propose a test statistic which is a nonparametric estimator of the moment condition $E[e_i E(e_i|X_i)p(X_i)]$. More specifically, the product $E(e_i|X_i)p(X_i)$ is estimated using the following kernel method

$$\widehat{E}(e_i|X_i)\widehat{p}(X_i) = \frac{1}{nh^m} \sum_{j=1}^{n} K\left(\frac{X_i - X_j}{h}\right) e_j,$$

where $K(\cdot)$ is a nonnegative kernel function and h is a smoothing parameter that converges to zero as the sample size increases. The outer-layer expectation of $E[e_i E(e_i|X_i)p(X_i)]$ is estimated by the sample average. Since the e_i's are not observable, they are replaced by $\widehat{e}_i = I(Y_i \ge g(X_i, \widehat{\beta})) - (1 - \tau)\widehat{G}(g(X_i, \widehat{\beta}))$, $i = 1, \ldots, n$, where $\widehat{\beta}$ is a $\sqrt{n}$-consistent estimator of β_0 and $\widehat{G}$ is the Kaplan-Meier estimator for G, the survival function of the censoring variable. This leads to our test statistic

$$T_n = \frac{1}{n(n-1)h^m} \sum_{i=1}^{n} \sum_{j \ne i} K\left(\frac{X_i - X_j}{h}\right) \widehat{e}_i \widehat{e}_j. \tag{5}$$

Large value of T_n gives evidence against the null hypothesis. Our proposed test thus generalizes the conditional moment test of Zheng (1998) for quantile regression without censoring.

There exists no uniform guidance on how to choose h for hypothesis testing. For the complete data case, Zheng (1998) proposed to apply generalized cross-validation (GCV) to choose h. For the random censoring situation, our simulations (see Section 4) suggest that this also often yields a reasonable choice. For median regression under random censoring, the GCV approach chooses h to minimize

$$GCV(h) = \frac{\sum_{i=1}^{n}[I(Y_i \geq X_i\widehat{\beta}) - 0.5\widehat{G}(g(X_i\widehat{\beta})) - \widehat{g}_h(x_i)]^2}{n[1 - tr(H)/n]^2}, \tag{6}$$

where

$$\widehat{g}_h(x) = \frac{\sum_{j=1}^{n} K((x - X_j)/h)[I(Y_i \geq X_i\widehat{\beta}) - 0.5\widehat{G}(g(X_i\widehat{\beta}))]}{\sum_{j=1}^{n} K((x - X_j)/h)}$$

and $tr(H)$ is the trace of the matrix $H = (h_{i,j})_{n\times n}$ with $h_{i,j} = K((X_i - X_j)/h)/\sum_{l=1}^{n} K((X_i - X_l)/h)$. Minimizing GCV generally yields an optimal bandwidth for estimation that is proportional to $n^{-1/(m+4)}$.

3. ASYMPTOTIC PROPERTIES

This section presents the asymptotic distribution of the proposed test under the null hypothesis and local alternative sequence. We first outline the regularity conditions used to establish the large sample theory.

Assumption 1. The observations $\{(Y_i, X_i, \Delta_i), i = 1, \ldots, n\}$ constitute a random sample for which $Y_i = min(T_i, C_i)$, where the T_i's are generated by (1), the C_i's are independent and identically distributed with cumulative distribution function $1 - G(\cdot)$, and the C_i's are also independent with the X_i's and T_i's.

Assumption 2. The probability density function $p(x)$ of X_i and its first-order derivative are uniformly bounded. In (1), the random errors ϵ_i's are independent and $P(\epsilon_i \leq 0|X_i) = \tau$. The ϵ_i's have bounded probability density functions.

Assumption 3. The parameter space B is a compact and convex subset of a Euclidean space. The true parameter value β_0 is an interior points in B. There exists an estimator $\widehat{\beta}_n$ such that under the null hypothesis $\sqrt{n}\left(\widehat{\beta}_n - \beta_0\right) = O_p(1)$.

Assumption 4. The quantile regression function $g(x, \beta)$ has second-order continuous derivative in β.

Assumption 5. The kernel function $K(u)$ is a symmetric density function with compact support and a continuous derivative.

The censoring times are assumed to be independent of the covariates and the random errors. This assumption is quite standard, although it might be restrictive in some applications. The same censoring scheme is also adopted by Ying et. al. (1995) for their main results (they discussed a heuristic discretization method for covariate dependent censoring) and Honoré et al. (2002), among others. Relaxation of this assumption is possible, but we will not exploit it in this paper. The validity of the proposed test only requires a $\sqrt{n}$-consistent estimator of β. The high-level conditions in Assumption 3 can be replaces by a set of more primitive conditions on the error distributions and the covariates, such as those listed in Honoré et al. (2002), which ensure the identifiability of the parameters (in particular their condition RC) and guarantee the $\sqrt{n}$-consistency of their estimator.

Under H_0, it can be shown that if we replace the unobservable e_i's in T_n by the estimators $\widehat{e}_i$'s, it will not change the asymptotic distribution of T_n. This result is stated in the following lemma.

LEMMA 1. Given assumptions 1-5, if $h \to 0$ and $n^{\alpha}h^{2m} \to \infty$ for some $0 < \alpha < 1$ as $n \to \infty$, then under H_0,

$$nh^{m/2}(T_n - T_n^*) \to 0$$

in probability as $n \to \infty$, where

$$T_n^* = \frac{1}{n(n-1)h^m} \sum_{i=1}^{n} \sum_{j \neq i} K\left(\frac{X_i - X_j}{h}\right) e_i e_j.$$

There are two technical complications for deriving the above results. One is due to the fact that $I(Y_i \geq g(X_i, \beta))$ is not differentiable in β, the other is that the survival function of the censoring variable $G(\cdot)$ is estimated by $\widehat{G}(\cdot)$ and its expansion leads to third-order U-statistics. To overcome these difficulties, we apply the uniform convergence results of Sherman (1994) for U-process indexed by parameters combined with a useful result of Zheng (1998) and Hoeffding decomposition of higher-order U-statistics. The details of the technical arguments of this lemma and the main theorems in this section are given in the appendix.

As a result of Lemma 1, the asymptotic distribution of interest is determined by $nh^{m/2}T_n^*$. Since $nh^{m/2}T_n^*$ has the form of a degenerate U-statistic and the e_i's are independent, the central limit theorem for degenerate U-statistic in Hall (1984) can then be applied to obtain the asymptotic normal distribution. This leads to

Theorem 1. Assume the conditions of Lemma 1, under H_0,

$$nh^{m/2}T_n \to N(0, \xi^2)$$

in distribution as $n \to \infty$, where

$$\xi^2 = 2\int K^2(u)du \int [\sigma^2(x)]^2 p^2(x)dx,$$

with $\sigma^2(x) = (1-\tau)G(g(x, \beta_0))[1 - (1-\tau)G(g(x, \beta_0))]$.

An asymptotic test thus can be based on the normal distribution in Theorem 1. The asymptotic variance ξ^2 can be consistently estimated by

$$\widehat{\xi^2} = \frac{2}{n(n-1)h^m} \sum_{i=1}^{n} \sum_{j \neq i} K^2\left(\frac{X_i - X_j}{h}\right) \widehat{e}_i^2 \widehat{e}_j^2.$$

For nonparametric tests like this, the power property is often investigated via considering a local alternative sequence that converges to the null at an appropriate rate, say d_n. If the local alternative sequence converges to H_0 faster than d_n, the test will not be able to distinguish the alternative from the null; on the other hand, if it converges to H_0 more slowly than d_n, the test will have asymptotic power equal to one. Thus d_n represents the rate that leads to a nondegenerate power of interest. In our setting of testing the fit of a conditional quantile function, we consider the following local alternative sequence

$$H_{1n} : F(g(X_i, \beta_0) + d_n l(X_i)|X_i) = \tau, \tag{7}$$

where $F(\cdot|X_i)$ is the conditional distribution function of T_i given X_i, $d_n = n^{-1/2}h^{-m/4}$ and $l(\cdot)$ is a function with continuous derivative. The following theorem gives the asymptotic distribution of the test statistic T_n under this local alternative sequence.

Theorem 2. Assume the conditions of Lemma 1, then under the local alternative sequence H_{1n} defined in (7),

$$nh^{m/2}T_n \to N(\eta, \xi^2)$$

in distribution as $n \to \infty$, where ξ^2 is defined in Theorem 1 and

$$\eta = E[l^2(X_i)f^2(g(X_i,\beta_0))G^2(g(X_i,\beta_0))p(X_i)],$$

with $f(\cdot)$ denoting the conditional density function of T_i given X_i.

The rate of local alternative $n^{-1/2}h^{-m/4}$ is the same as that of kernel smoothing test for checking the lack-of-fit of a mean regression model or that of a quantile regression model without censoring. By letting h converge to zero slowly enough, this rate can be made arbitrarily close to the parametric rate $n^{-1/2}$.

4. SIMULATION RESULTS

In this section, we report results from several Monte Carlo studies to demonstrate the performance of the proposed test for finite sample size. The critical value for the test is taken from the asymptotic normal distribution in Theorem 1. The simulations are based on 500 runs with specified level 0.05. For these designs, the overall censoring probabilities are between 40% and 50% under the null hypothesis; and vary with the specific alternative. For example, the quadratic alternative in Example 2 corresponds to a censoring rate approximately between 30%-50% for the different error distributions under consideration.

Example 1. We first consider an example of testing the linearity of the conditional median function with a single covariate when the response is under right random censoring. This, in particular, allows for a straightforward comparison with a parametric test based on the bootstrap confidence interval (abbreviated as BT test in the sequel), where an alternative model needs to be explicitly specified. The BT test rejects the null hypothesis if the bootstrapped confidence interval for the coefficient of the quadratic term of the covariate does not contain zero. The estimation of β and the BT test are implemented by the "Censored Regression Quantiles" library of the software package R (see Portnoy, 2003); and the smoothing test uses the Epanechnikov kernel $K(u) = 0.75(1-u^2)I(|u| \leq 1)$.

Under the null hypothesis, the survival time T_i is generated by

$$T_i = -0.7 + X_i + \epsilon_i, \quad i = 1, \ldots, 100, \tag{8}$$

where the X_i's are uniformly distributed on [0,1], the censoring variable is uniformly distributed on [-1.5,1.5], and the random error has a standard normal distribution. To investigate the power performance, we consider three different alternatives. The simulation setting is the same as above except that now

$$T_i = -0.7 + X_i + \epsilon_i - al(X_i), \tag{9}$$

where $a = 1, 2, \ldots, 10$ and $l(X_i)$ takes one of the following forms: (1) exponential alternative $l(X_i) = 0.05exp(X_i)$, (2) quadratic alternative $l(X_i) = (X_i - 0.5)^2$, and (3) cosine alternative $l(X_i) = 0.1cos(X_i)$.

To investigate the effect of bandwidth h, we consider $h = cn^{-1/5}$, for $c = 0.5, 1.0, 1.5, 2.0, 2.5$, which satisfies the requirement of the theorem. For the data generated under the null hypothesis, the T_n test gives estimated levels 0.044, 0.040, 0.048, 0.058 and 0.068 corresponding to the different choices of c, and the BT test gives an estimated level of 0.062. Both tests have observed levels close to the specified nominal level 0.05. For each of the three alternatives, we estimate the power of both the T_n test (for different bandwidth) and the BT test for different values of a. The estimated power curves for each of the three alternatives are depicted in Figures 1-3. For the quadratic alternative, the BT test and the T_n test behave similarly. But for the exponential alternative and the cosine alternative, the BT test has very low power (close to its nominal level) while the T_n test

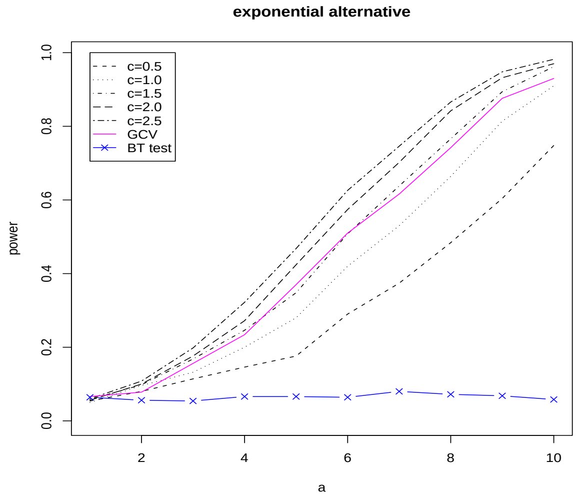

Figure 1: Simulated power curves for exponential alternative.

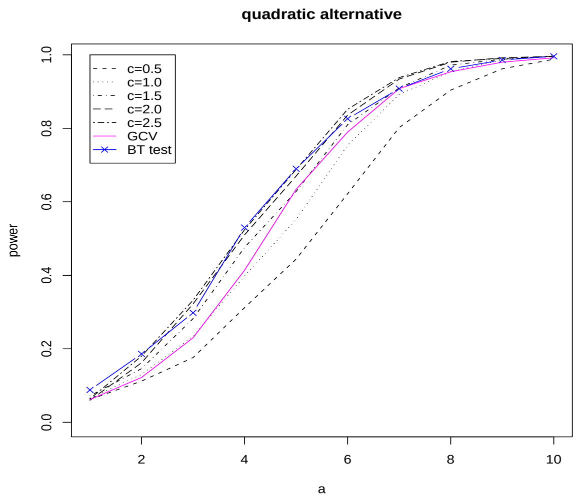

Figure 2: Simulated power curves for quadratic alternative.

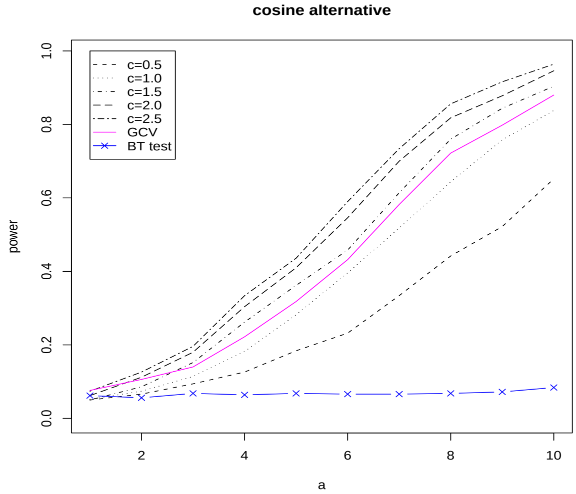

Figure 3: Simulated power curves for cosine alternative.

displays much higher power. In general, tests that require to specify a parametric alternative may suffer from low power or break down if the specified alternative is different from the true model.

The power of T_n is sensitive to the choice of the smoothing parameter h. We next examine the effectiveness of the GCV approach to choose h. More specifically, for this example we set the bandwidth to be $\lambda n^{-1/5}$ and use grid search to find the optimal λ that minimizes the GCV criterion (6) for $0.5 \le \lambda \le 2.5$, where the grid point has increment 0.1. For the data generated under the null hypothesis, the GCV approach yields an estimated level of 0.052; for the above three alternatives, the corresponding power curves are depicted by the solid line in Figures 1-3. Although in general GCV does not yield optimal power, its performance is quite satisfactory.

Example 2. Next, we further study the effectiveness of the GCV approach in a two-dimensional setting, where the bivariate Epanechnikov kernel $K(u_1, u_2) = K(u_1)K(u_2)$ with $K(u) = 0.75(1 - u^2)I(|u| \le 1)$ is used. The survival time T_i is generated by

$$T_i = -1.2 + X_{1i} + X_{2i} - l(X_{1i}, X_{2i}) + \epsilon_i, \quad i = 1, \ldots, 100,$$

where the X_{1i}'s and X_{2i}'s are independently uniformly distributed on [0,1], and the censoring variable has a standard normal distribution. We consider four different error distributions: standard normal distribution, t distribution with 3 degrees of freedom, standard lognormal distribution subtracted by one, which gives median zero, and heteroscedastic errors $\epsilon_i = exp(0.5x_{1i})e_i$, where the e_i's are independent standard normal random variables. We also consider four different functional forms for $l(X_{1i}, X_{2i})$: (1) $l(X_{1i}, X_{2i}) = 0$, corresponding to the null hypothesis; (2) $l(X_{1i}, X_{2i}) = 0.5(X_{1i}^2 + X_{1i}X_{2i} + X_{2i}^2)$; (3) $l(X_{1i}, X_{2i}) = 2cos((X_{1i} + X_{2i}))$ and (4) $l(X_{1i}, X_{2i}) = 0.5exp(0.5X_{1i})$. Table 1 reports the results of the smoothing test T_n with the smoothing parameter chosen via GCV, which has satisfactory performance under both the null and the alternatives. In the bivariate case, it becomes more challenging to apply the BT test since unlike the single covariate case, a natural parametric alternative is not immediately available thus we have not reported the BT test in this example .

5. STANFORD HEART TRANSPLANT DATA ANALYSIS

Table 1: The proportion of times the null hypothesis of linearity is rejected by the T_n test and the test based on the bootstrap confidence intervals.

$l(X_1, X_2)$	$n(0,1)$	t_3	lognormal	heteroscedastic
(1)	0.050	0.046	0.020	0.060
(2)	0.612	0.600	0.568	0.640
(3)	0.854	0.796	0.566	0.850
(4)	0.780	0.770	0.644	0.806

To illustrate the application, we apply the proposed test to the well known Stanford Heart Transplant data. This data set was originally reported in Miller and Halpern (1982) and has been analyzed by many authors in the literature, including Honoré, Khan and Powell (2002). The response variable is the survival time of patients who had received heart transplants through the Stanford heart transplant program between October 1967 and February 1980. The covariate of interest is age at the time of the first transplant. A censoring indicator equals one if the patient was dead (uncensored) and zero (censored) otherwise. Following the literature, we include in our analysis 157 patients whose tissue mismatch scores were not missing and take the response variable as the logarithm (base 10) of the observed survival time. We use the "Censored Regression Quantiles" library of R to estimate β and to perform the BT test. The smoothing test uses the bivariate Epanechnikov kernel.

Honoré, Khan and Powell (2002) analyzed this data set using censored quantile regression and suggests that age has a significant quadratic effect on the logarithm of the survival time of the heart transplant patients. We will show below that our model checking technique can be used to verify their model.

We consider median regression model under random censoring. We first test the null hypothesis that the conditional median of the logarithm of the survival time given age can be adequately described by a linear function of age. For applying a smoothing test to a real data set, King, Hart and Wehrly (1991), Young and Bowman (1995) suggest calculating the p-value of the test for different choices of the smoothing parameter, and called the plot of p-value versus the smoothing parameter h a "significant trace". Any reasonable smoothing test should either reject or not reject the null hypothesis for a wide range choices of smoothing parameter. We adopt this suggestion here. We standardize the covariate age to the interval [0,1]. For a wide range choice of smoothing parameter h between 0 and 1, the p-value of the T_n test is almost zero. The GCV approach chooses $h = 0.255$. This strongly indicates that a linear conditional median function is not supported by the data. The BT test also rejects the null hypothesis when testing linear age effect versus quadratic age effect.

We next test the null hypothesis that the conditional median of the logarithm of the survival time given age can be adequately described by a quadratic function of age. Except for very small bandwidth (close to zero) and very large bandwidth (close to one), the p-value of the test is well above 0.05. The plot of the "significance trace" for testing this hypothesis is given in Figure 4. For this hypothesis, the GCV approach chooses $h = 0.258$. The BT test yields the same conclusion when testing quadratic age effect versus cubic age effect. This suggests that a quadratic function of age provides a reasonable fit of the conditional median of the logarithm of the survival time given age.

6. DISCUSSIONS

The results in the paper pertain to continuous covariates case. As in Zheng (1998), the theory

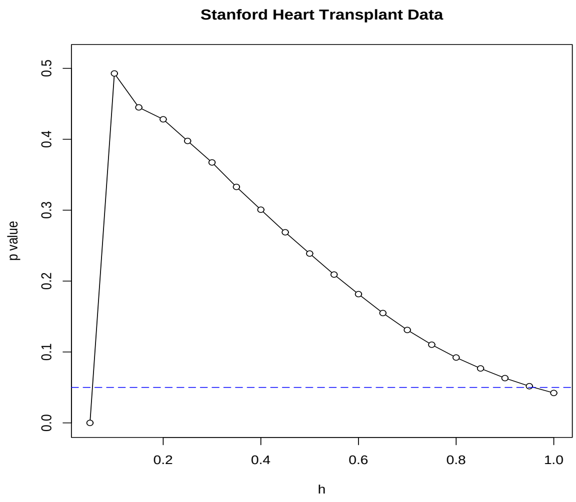

Figure 4: Significance trace for testing the quadratic effect of age for a wide range choice of the smoothing parameter h. The dashes horizontal line corresponds to nominal level of the test: 0.05.

may be extended to include discrete covariates. To provide a basic idea, write $X_i = (X_{1i}', X_{2i}')$, where X_{1i} is an m_1-dimensional vector of discrete covariates, X_{2i} is an m_2-dimensional vector of continuous covariates, and $m_1 + m_2 = m$. Then the test statistic (5) now has the form

$$T_n = \frac{1}{n(n-1)h^{m_2}} \sum_{i=1}^{n} \sum_{j \neq i} K\left(\frac{X_i - X_j}{h}\right) \widehat{e}_i \widehat{e}_j.$$

Note that $h \to 0$ when $n \to \infty$. When n is large, only observations from the same group (corresponding to different levels of the discrete covariates) will be treated as local. The asymptotic analysis can still be similarly carried out. We have not pursued this direction since in practice the mixed covariates case is more like analysis of covariance, where the discrete covariates are treated as factors, and model checking mainly focuses on the functional form of the continuous covariates. In this situation, it might be more efficient to consider a more structured semiparametric model such as the censored partial linear quantile regression model (Chen and Khan, 2001).

As suggested by an anonymous referee, we may obtain a simple visual diagnostic by implementing the recently proposed local linear estimator for censored quantile regression (Gannoun et al., 2005) and observing whether it agrees with the parametric fit.

Just as nonparametric smoothing estimators, the proposed nonparametric smoothing test suffers from the curse of dimensionality. However, the parametric tests avoid this only by imposing strong structural assumptions, which are hard to verify in high dimension. They pay the possible price of serious power loss when the parametric assumptions are violated. On the other hand, parametric tests, when all related parametric assumptions are satisfied, may have superior power property. An alternative approach for model checking with nonparametric alternative in this setting is to generalize the empirical process type test of He and Zhu (2003).

APPENDIX

The proofs use uniform convergence result for U-statistic indexed by parameters. Following the notations of Sherman (1994), the second-order U-statistic indexed by β is defined as

$$U_n^2 H_n(\cdot, \beta) = \frac{1}{n(n-1)} \sum_{i \neq j} H_n(Z_i, Z_j, \beta),$$

where $\sum_{i\neq j}$ denotes the double sum $\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}$, and the function $H_n(Z_i, Z_j, \beta)$ needs not be symmetric in Z_i and Z_j. The following lemma is useful:

Lemma 1 of Zheng (1998). Let $H = \{H_n(\cdot,\beta), \beta \in B\}$ be a class of degenerate functions, i.e., $E(H_n(Z_i, Z_j, \beta)|Z_i) = E(H_n(Z_i, Z_j, \beta)|Z_j)$=0. If (i) H is Euclidean for a constant envelope function and (ii) $E[H_n(\cdot,\beta) - H_n(\cdot,\beta_0)]^2 \leq O(n^{-1/2})$ uniformly over $O_p(n^{-1/2})$ neighborhood of β_0, then for any $0 < \alpha < 1$,

$$\left|U_n^2 H_n(\cdot,\beta) - U_n^2 H_n(\cdot,\beta_0)\right| \leq O_p(n^{-1-\alpha/4})$$

uniformly over $O_p(n^{-1/2})$ neighborhood of β_0.

Proof of Lemma 1. Note that

$$\widehat{e}_i - e_i = [I(Y_i \geq g(X_i,\widehat{\beta})) - I(Y_i \geq g(X_i,\beta_0))] - (1-\tau)[\widehat{G}(g(X_i,\widehat{\beta})) - G(g(X_i,\beta_0))].$$

Therefore

$$\begin{aligned}
T_n &= T_n^* + \frac{2}{n(n-1)h^m}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right) e_i[I(Y_j \geq g(X_j,\widehat{\beta})) - I(Y_j \geq g(X_j,\beta_0))] \\
&\quad - \frac{2(1-\tau)}{n(n-1)h^m}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right) e_i[\widehat{G}(g(X_j,\widehat{\beta})) - G(g(X_j,\beta_0))] \\
&\quad + \frac{1}{n(n-1)h^m}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right)[I(Y_i \geq g(X_i,\widehat{\beta})) - I(Y_i \geq g(X_i,\beta_0))] \\
&\qquad\qquad \times[I(Y_j \geq g(X_j,\widehat{\beta})) - I(Y_j \geq g(X_j,\beta_0))] \\
&\quad - \frac{2(1-\tau)}{n(n-1)h^m}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right)[I(Y_i \geq g(X_i,\widehat{\beta})) - I(Y_i \geq g(X_i,\beta_0))] \\
&\qquad\qquad \times[\widehat{G}(g(X_j,\widehat{\beta})) - G(g(X_j,\beta_0))] \\
&\quad + \frac{(1-\tau)^2}{n(n-1)h^m}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right)[\widehat{G}(g(X_i,\widehat{\beta})) - G(g(X_i,\beta_0))] \\
&\qquad\qquad \times[\widehat{G}(g(X_j,\widehat{\beta})) - G(g(X_j,\beta_0))] \\
&= T_n^* + \sum_{t=1}^{5} D_t,
\end{aligned}$$

where the definition of D_t should be clear from the context. To prove the lemma, we need to establish $nh^{m/2}D_t = o_p(1)$, $t = 1, \ldots, 5$. We first verify

$$\begin{aligned}
nh^{m/2}D_1 &= \frac{2}{(n-1)h^{m/2}}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right) e_i[I(Y_j \geq g(X_j,\widehat{\beta})) - I(Y_j \geq g(X_j,\beta_0))] \\
&= o_p(1).
\end{aligned}$$

It is sufficient to show $sup_{||\beta-\beta_0||\leq cn^{-1/2}}|D_1(\beta,\beta_0)| = o_p(1)$ for any positive constant c, where $D_1(\beta,\beta_0) = 2[(n-1)h^{m/2}]^{-1}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right) e_i[I(Y_j \geq g(X_j,\beta)) - I(Y_j \geq g(X_j,\beta_0))]$. Let $H(\cdot)$ be the conditional CDF of Y_i given X_i, then

$$\begin{aligned}
& D_1(\beta,\beta_0) \\
=\ & \frac{2}{(n-1)h^{m/2}}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right) e_i[I(Y_j \geq g(X_j,\beta)) - I(Y_j \geq g(X_j,\beta_0)) \\
& \qquad - H(g(X_j,\beta)) + H(g(X_j,\beta_0))]
\end{aligned}$$

$$
\begin{aligned}
&+\frac{2}{(n-1)h^{m/2}}\sum_{j\neq i}K\left(\frac{X_i-X_j}{h}\right)e_i[H(g(X_j,\beta))-H(g(X_j,\beta_0))]\\
=\ & D_{11}(\beta,\beta_0)+D_{12}(\beta,\beta_0),
\end{aligned}
$$

where the definitions of $D_{11}(\beta,\beta_0)$ and $D_{12}(\beta,\beta_0)$ should be clear from the context. Write $Q_{11}(\beta,\beta_0)=(2n)^{-1}h^{m/2}D_{11}(\beta,\beta_0)=[n(n-1)]^{-1}\sum_{j\neq i}H_{1n}(Z_i,Z_j,\beta)$, where $Z_i=(X_i,Y_i)$ and $H_{1n}(Z_i,Z_j,\beta)=K\left(\frac{X_i-X_j}{h}\right)e_i[I(Y_j\geq g(X_j,\beta))-I(Y_j\geq g(X_j,\beta_0))-H(g(X_j,\beta))+H(g(X_j,\beta_0))]$, then $Q_{11}(\beta,\beta_0)$ is a second-order degenerate U-statistic. If we let $F_{1n}=\{H_{1n}(Z_i,Z_j,\beta),\beta\in B\}$, then F_{1n} is a Euclidean class of functions with a constant envelope function. Moreover,

$$
\begin{aligned}
& E[H_{1n}(Z_i,Z_j,\beta)-H_{1n}(Z_i,Z_j,\beta_0)]^2\\
=\ & E\left\{K^2\left(\frac{X_i-X_j}{h}\right)e_i^2[I(Y_j\geq g(X_j,\beta))-I(Y_j\geq g(X_j,\beta_0))-H(g(X_j,\beta))+H(g(X_j,\beta_0))]^2\right\}\\
\leq\ & CE\left\{[I(Y_j\geq g(X_j,\beta))-I(Y_j\geq g(X_j,\beta_0))-H(g(X_j,\beta))+H(g(X_j,\beta_0))]^2\right\}\\
\leq\ & CE\left\{E\{[I(Y_j\geq g(X_j,\beta))-I(Y_j\geq g(X_j,\beta_0))]^2|X_j\}\right\}\\
=\ & CE[H(g(X_j,\beta))+H(g(X_j,\beta_0))-2H(min(g(X_j,\beta),g(X_j,\beta_0)))]\\
\leq\ & C||\beta-\beta_0||,
\end{aligned}
$$

where C is some positive constant. Thus by the lemma of Zheng (1998),

$$
sup_{||\beta-\beta_0||\leq cn^{-1/2}}|Q_{11}(\beta,\beta_0)|\leq O_p(n^{-1-\alpha/4}),\quad \forall\ 0<\alpha<1.
$$

So $sup_{||\beta-\beta_0||\leq cn^{-1/2}}|D_1(\beta,\beta_0)|=O(nh^{-m/2})O_p(n^{-1-\alpha/4})=O_p(h^{-m/2}n^{-\alpha/4})=o_p(1)$. Note that

$$
\begin{aligned}
D_{12}(\beta,\beta_0) =\ & \frac{2}{(n-1)h^{m/2}}\sum_{j\neq i}K\left(\frac{X_i-X_j}{h}\right)e_i(\beta-\beta_0)'h(X_j,\beta_0)\\
& +\frac{1}{(n-1)h^{m/2}}\sum_{j\neq i}K\left(\frac{X_i-X_j}{h}\right)e_i(\beta-\beta_0)'T_j(\tilde\beta)(\beta-\beta_0)\\
=\ & \frac{2}{(n-1)h^{m/2}}(\beta-\beta_0)'Q_{12}^A(\beta_0)+\frac{n}{h^{m/2}}(\beta-\beta_0)'Q_{12}^B(\tilde\beta,\beta_0)(\beta-\beta_0),
\end{aligned}
$$

where $h(X_j,\beta_0)=\frac{\partial H(g(X_j,\beta))}{\partial\beta}|_{\beta=\beta_0}$, $T_j(\beta)=\frac{\partial^2H(g(X_j,\beta))}{\partial\beta\partial\beta'}$, $\tilde\beta$ is between β_0 and β, and the definition of $Q_{12}^A(\beta_0)$ and $Q_{12}^B(\tilde\beta,\beta_0)$ should be clear from the context. The order of $Q_{12}^A(\beta_0)$ could be determined by checking its mean and variance. We have $E(Q_{12}^A(\beta_0))=0$ and

$$
\begin{aligned}
E[(Q_{12}^A(\beta_0))^2] =\ & E\left[\sum_{j_1\neq i_1}\sum_{j_2\neq i_2}K\left(\frac{X_{i1}-X_{j1}}{h}\right)K\left(\frac{X_{i2}-X_{j2}}{h}\right)e_{i_1}e_{i_2}h(X_{j1},\beta_0)h(X_{j2},\beta_0)\right]\\
=\ & O(n^3h^{2m}).
\end{aligned}
$$

Thus $Q_{12}^A(\beta_0)=O_p(n^{3/2}h^m)$. For $Q_{12}^B(\beta,\beta_0)$, we have that for some positive constant C,

$$
E\left[sup_{||\beta-\beta_0||\leq cn^{-1/2}}|Q_{12}^B(\beta,\beta_0)|\right]\leq CE\left[\frac{1}{n(n-1)}\sum_{j\neq i}K\left(\frac{X_i-X_j}{h}\right)\right]=O(h^m).
$$

Thus $sup_{||\beta-\beta_0||\leq cn^{-1/2}}|D_{12}((\beta,\beta_0))|=O_p(h^{m/2})=o_p(1)$. By now, we've proved $nh^{m/2}D_1=$

$o_p(1)$. Next, we will show $nh^{m/2}D_2 = o_p(1)$. Note

$$\begin{aligned} nh^{m/2}D_2 &= -\frac{1}{(n-1)h^{m/2}}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right) e_i[\widehat{G}(g(X_j,\beta_0)) - G(g(X_j,\beta_0))] \\ &\quad -\frac{1}{(n-1)h^{m/2}}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right) e_i[G(g(X_j,\widehat{\beta})) - G(g(X_j,\beta_0))] \\ &\quad -\frac{1}{(n-1)h^{m/2}}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right) e_i[\widehat{G}(g(X_j,\widehat{\beta})) - G(g(X_j,\widehat{\beta})) \\ &\qquad\qquad -\widehat{G}(g(X_j,\beta_0)) + G(g(X_j,\beta_0))] \\ &= D_{21} + D_{22} + D_{23}, \end{aligned}$$

where the definition of D_{2t}, $t = 1, 2, 3$, should be clear from the context. Similarly as above, we can show that $D_{22} = o_p(1)$ by proving that for any $c > 0$,

$$sup_{||\beta-\beta_0||\leq cn^{-1/2}} \left| \frac{1}{(n-1)h^{m/2}}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right) e_i[G(g(X_j,\beta)) - G(g(X_j,\beta_0))] \right| = o_p(1).$$

To evaluate the order of D_{21}, we use the almost surely iid representation of the Kaplan-Meier estimator by Lo and Singh (1986) and obtain

$$\begin{aligned} D_{21} &= -\frac{1}{(n-1)h^{m/2}}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right) e_i \left[\frac{1}{n}\sum_{l=1}^{n} q(Y_l, \Delta_l, g(X_j,\beta_0))\right] \\ &\quad -\frac{1}{(n-1)h^{m/2}}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right) e_i r_n(g(X_j,\beta_0)) \\ &= D_{211} + D_{212}, \end{aligned}$$

where the definition of D_{211} and D_{212} should be clear from the context; the notation $q(Y_l, \Delta_l, g(X_j,\beta_0)))$ and $r_n(g(X_j,\beta_0))$ can be found in Lo and Singh (1986), and $r_n(\cdot)$ has the property that $sup_{0\leq t\leq T}|r_n(t)| = O(n^{-3/4}(logn)^{3/4})$ almost surely where T is the upper bound of t. We thus have

$$\begin{aligned} |D_{212}| &\leq O\left(\frac{1}{nh^{m/2}}\right)\sum_{j=1}^{n}\left|\sum_{\substack{i=1\\ i\neq j}}^{n} K\left(\frac{X_i - X_j}{h}\right) e_i\right| |r_n(g(X_j,\beta_0))| \\ &\leq O\left(\frac{1}{nh^{m/2}}\right) O(n) O_p(n^{1/2}h^{m/2}) O(n^{-3/4}(logn)^{3/4}) \\ &= O_p(n^{-1/4}(logn)^{3/4}) = o_p(1). \end{aligned}$$

Thus to show $D_{21} = o_p(1)$, it is sufficient to show $D_{211} = o_p(1)$. Note

$$D_{211} = -\frac{1}{n(n-1)h^{m/2}}\sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j\neq l}}^{n}\sum_{l=1}^{n} K\left(\frac{X_i - X_j}{h}\right) e_i q(Y_l, \Delta_l, g(X_j,\beta_0)) + o_p(1).$$

Denote the first term of the right side of the above expression by D^*_{211}. The result of Lo and Singh (1986) indicates that conditional on the covariates, $q(Y_l, \Delta_l, g(X_j,\beta_0)))$, $i = 1,\ldots,n$, are independent with mean zero. By checking mean and variance, it is easy to verify that $D^*_{211} = o_p(1)$.

It remains to show $D_{23} = o_p(1)$. By Lo and Singh's iid representation of the Kaplan-Meier estimator, it is sufficient to show that for

$$\begin{aligned} & D_{23}^*(\beta, \beta_0) \\ = & -\frac{1}{n(n-1)h^{m/2}} \sum_{i=1}^n \sum_{\substack{j=1 \\ i\neq j}}^n \sum_{l=1}^n K\left(\frac{X_i - X_j}{h}\right) e_i[q(Y_l, \Delta_l, g(X_j, \beta)) - q(Y_l, \Delta_l, g(X_j, \beta_0))], \end{aligned}$$

we have $sup_{||\beta-\beta_0|| \leq cn^{-1/2}} |D_{23}^*(\beta, \beta_0)| = o_p(1)$ for any positive constant c. Note

$$\begin{aligned} & sup_{||\beta-\beta_0|| \leq cn^{-1/2}} |D_{23}^*((\beta, \beta_0))| \\ \leq & \frac{1}{n(n-1)h^{m/2}} \left| \sum_{i=1}^n \sum_{\substack{j=1 \\ i\neq j\neq l}}^n \sum_{l=1}^n K\left(\frac{X_i - X_j}{h}\right) e_i[q(Y_l, \Delta_l, g(X_j, \beta)) - q(Y_l, \Delta_l, g(X_j, \beta_0))] \right| \\ & + o_p(1). \end{aligned}$$

Let $h(Z_i, Z_j, Z_l, \beta) = K\left(\frac{X_i - X_j}{h}\right) e_i[q(Y_l, \Delta_l, g(X_j, \beta)) - q(Y_l, \Delta_l, g(X_j, \beta_0))]$, then we need to show $(n-2)h^{-m/2} sup_{||\beta-\beta_0|| \leq cn^{-1/2}} |U_n^3 h(Z_i, Z_j, Z_l, \beta)| = o_p(1)$ for any positive constant c, where $U_n^3 h(Z_i, Z_j, Z_l, \beta)$ denotes a third-order U-statistic with kernel function $h(Z_i, Z_j, Z_l, \beta)$, see the definition in Sherman (1994). Using Hoeffding's method of decomposition (see Section 11.4 of van der Vaart, 1998), $U_n^3 h(Z_i, Z_j, Z_l, \beta)$ can be written as sum of degenerate U-statistics. In the notation of Sherman (1994, page 449),

$$U_n^3 h(Z_i, Z_j, Z_l, \beta) = P^3 h(Z_i, Z_j, Z_l, \beta) + P_n h_1(Z_i, Z_j, Z_l, \beta) + \sum_{i=2}^3 U_n^i h_i(Z_i, Z_j, Z_l, \beta).$$

Since the four expectations $E(h(Z_i, Z_j, Z_l, \beta))$,$E(h(Z_i, Z_j, Z_l, \beta)|Z_i)$, $E(h(Z_i, Z_j, Z_l, \beta)|Z_j)$ and $= E(h(Z_i, Z_j, Z_l, \beta)|Z_l)$ are all equal to zero, we have $P^3 h(Z_i, Z_j, Z_l, \beta) = P_n h_1(Z_i, Z_j, Z_l, \beta)$ $= 0$. $U_n^3 h(Z_i, Z_j, Z_l, \beta)$ thus can be expressed as sum of second-order degenerate U-statistic and third-order degenerate U-statistic. Therefore,

$$\begin{aligned} & sup_{||\beta-\beta_0|| \leq cn^{-1/2}} |U_n^3 h(Z_i, Z_j, Z_l, \beta)| \\ \leq & sup_{||\beta-\beta_0|| \leq cn^{-1/2}} |U_n^2 h_2(Z_i, Z_j, Z_l, \beta)| + sup_{||\beta-\beta_0|| \leq cn^{-1/2}} |U_n^3 h_3(Z_i, Z_j, Z_l, \beta)|. \end{aligned}$$

As in Lemma A.3 of Honore, Khan and Powell (2002), the class of functions $\{h(Z_i, Z_j, Z_l, \beta), \beta \in B\}$ is Euclidean with a constant envelope function. Then by Lemma 5 of Sherman (1994), the classes of functions $\{h_2(Z_i, Z_j, Z_l, \beta), \beta \in B\}$ and $\{h_3(Z_i, Z_j, Z_l, \beta), \beta \in B\}$ are both Euclidean with constant envelope functions. Thus, according to the lemma of Zheng,

$$\begin{aligned} (n-2)h^{-m/2} sup_{||\beta-\beta_0|| \leq cn^{-1/2}} |U_n^3 h(Z_i, Z_j, Z_l, \beta)| & \leq O(nh^{-m/2}) O_p(n^{-1-\alpha/4}) \\ & = O_p(h^{-m/2} n^{-\alpha/4}) = o_p(1). \end{aligned}$$

This finishes proving $nh^{m/2} D_2 = o_p(1)$. Similarly, we can show that $nh^{m/2} D_t = o_p(1)$ for $t = 3, 4, 5$.

Proof of Theorem 1. The conclusion of Lemma 1 indicates that $nh^{m/2} T_n$ and $nh^{m/2} T_n^*$ have the same asymptotic distribution. Since $nh^{m/2} T_n^*$ has the form of a second-order degenerate U-statistic, the central limit theorem of Hall (1984) can be directly applied to establish the asymptotic normality.

Proof of Theorem 2. Similarly as in the proof for Lemma 1, we can show that $nh^{m/2}T_n = nh^{m/2}T_n^* + o_p(1)$. We can write T_n^* as

$$\begin{aligned} T_n^* &= \frac{1}{n(n-1)h^m}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right)[I(Y_i \geq g(X_i,\beta)) - P_{1n}(Y_i \geq g(X_i,\beta_0)|X_i)] \\ &\qquad \times[I(Y_j \geq g(X_j,\beta)) - P_{1n}(Y_j \geq g(X_j,\beta_0)|X_i)] \\ &\quad + \frac{2}{n(n-1)h^m}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right)[P_{1n}(Y_i \geq g(X_i,\beta_0)|X_i) - (1-\tau)G(g(X_i,\beta_0))] \\ &\qquad \times[I(Y_j \geq g(X_j,\beta)) - P_{1n}(Y_j \geq g(X_j,\beta_0)|X_i)] \\ &\quad + \frac{1}{n(n-1)h^m}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right)[P_{1n}(Y_i \geq g(X_i,\beta_0)|X_i) - (1-\tau)G(g(X_i,\beta_0))] \\ &\qquad \times[P_{1n}(Y_j \geq g(X_j,\beta_0)|X_j) - (1-\tau)G(g(X_j,\beta_0))] \\ &= Q_1 + Q_2 + Q_3, \end{aligned}$$

where $P_{1n}(Y_i \geq g(X_i,\beta_0)|X_i)$ denotes the conditional probability under H_{1n}, and the definition of Q_t, $t = 1,2,3$, should be clear from the context. Same as in the proof of Theorem 1, we can show that $nh^{m/2}Q_1$ converges in distribution to the normal distribution under H_0. By Taylor expansion, under the local alternative H_{1n},

$$\begin{aligned} & P_{1n}(Y_i \geq g(X_i,\beta_0)|X_i) \\ =\ & [1 - F(g(X_i,\beta_0))]G(g(X_i,\beta_0)) \\ =\ & [1 - F(g(X_i,\beta_0) + d_n l(X_i)) + d_n l(X_i) f(g(X_i,\beta_0)) + O(d_n^2)]G(g(X_i,\beta_0)) \\ =\ & (1-\tau)G(g(X_i,\beta_0)) + d_n l(X_i) f(g(X_i,\beta_0))G(g(X_i,\beta_0)) + O(d_n^2). \end{aligned}$$

We therefore have for $d_n = n^{-1/2}h^{-m/4}$,

$$\begin{aligned} nh^{m/2}Q_2 &= \frac{2d_n}{(n-1)h^{m/2}}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right)[l(X_i)f(g(X_i,\beta_0))G(g(X_i,\beta_0)) + O(d_n)] \\ &\qquad \times[I(Y_j \geq g(X_j,\beta)) - P_{1n}(Y_j \geq g(X_j,\beta_0)|X_i)] \\ &= \frac{2d_n}{(n-1)h^{m/2}}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right) l(X_i)f(g(X_i,\beta_0))G(g(X_i,\beta_0)) \\ &\qquad \times[I(Y_j \geq g(X_j,\beta)) - P_{1n}(Y_j \geq g(X_j,\beta_0)|X_i)] + o_p(1) \\ &= Q_2^* + o_p(1). \end{aligned}$$

Direct calculation leads to $E(Q_2^*) = 0$ and $E((Q_2^*)^2) = O(n^{-3}h^{-3m/2})[O(n^3h^{2m}) + O(n^2h^m)] = o(1)$, thus $Q_2^* = o_p(1)$. Finally,

$$\begin{aligned} nh^{m/2}Q_3 &= \frac{d_n^2}{(n-1)h^{m/2}}\sum_{j\neq i} K\left(\frac{X_i - X_j}{h}\right) l(X_i)f(g(X_i,\beta_0))G(g(X_i,\beta_0)) \\ &\qquad \times l(X_j)f(g(X_j,\beta_0))G(g(X_j,\beta_0)) + o_p(1) \\ &= nh^{-m/2}d_n^2 \iint K\left(\frac{x-y}{h}\right) l(x)f(g(x,\beta_0))G(g(x,\beta_0)) \\ &\qquad \times l(y)f(g(y,\beta_0))G(g(y,\beta_0))p(x)p(y)dxdy + o_p(1) \\ &= nh^{m/2}d_n^2 \iint K(u)l(y+uh)f(g(y+uh,\beta_0))G(g(y+uh,\beta_0)) \\ &\qquad \times l(y)f(g(y,\beta_0))G(g(y,\beta_0))p(y+uh)p(y)dudy + o_p(1) \\ &= \iint K(u)l^2(y)f^2(g(y,\beta_0))G^2(g(y,\beta_0))p^2(y)dydu + o_p(1) = \eta + o_p(1), \end{aligned}$$

where η is defined in Theorem 2.

ACKNOWLEDGEMENTS

The author would like to thank the referees, the AE and the editor, whose comments significantly improve the paper. The author would also like to thank Ingrid van Keilegom for pointing out a useful result of Lo and Singh.

REFERENCES

S. Chen & S. Khan (2001). Semiparametric estimation of a partially linear censored regression model. *Econometric Theory,* 17, 567-590.

A. Gannoun, J. Saracco, A. Yuan & G. E. Bonney (2005). Non-parametric Quantile Regression with Censored Data. *Scandinavian Journal of Statistics,* 32, 527-550.

P. Hall (1984). Central limit theorem for integrated square error of multivariate nonparametric density estimators. *Journal of Multivariate Analysis,* 14, 1-16.

X. He & L. X. Zhu (2003). A lack-of-fit test for quantile regression. *Journal of the American Statistical Association, 98, 1013-1022.*

J. L. Horowitz & V. G. Spokony (2002). An adptive, rate-optimal test of linearity for median regression models. *Journal of the American Statistical Association,* 97 822-835.

B. Honoré, S. Khan & J. L. Powell (2002). Quantile regression under random censoring. *Journal of Econometrics,* 109 67-105.

J. Huang, S. G. Ma & H. L. Xie (2005). Least absolute deviations estimation for the accelerated failure time model. Technical report. Department of Statistics and Actuarial Science, University of Iowa.

E. King, J. D. Hart & T. E. Wehrly (1991). Testing the equality of two regression curves using linear smoothers. *Statistics & Probability Letters* 12, 239-247.

S. H. Lo & K. Singh (1986). The product-limit estimator and the bootstrap: some asymptotic representations. *Probability Theory and Related Fields,* 71 455-465.

R. Miller & J. Halpern (1982). Regression with censored data. *Biometrika,* 69 521-531.

S. Portnoy (2003). Censored regression quantiles. *Journal of the American Statistical Association,* 98 1001-1012.

J. L. Powell (1984). Least absolue deviations estimation for the censored regression model. *Journal of Econometrics*, 25, 303-325.

G. S. Qin & M. Tsao (2003). Empirical likelihood inference for median regression models of censored survival data. *Journal of Multivariate Analysis*, 85, 416-430.

R. S. Sherman (1994). Maximal inequalities for degenerate U-processes with applications to optimization estimators. *The Annals of Statistics,* 22 439-459.

S. Subramanian (2002). Median regression using nonparametric kernel estimation. *Journal of Nonparametric Statistics*, 14, 583-605.

M. Tableman & J. S. Kim (2000). Survival analysis using S: Analysis of time-to-event data. Chapman and Hall/CRC.

A. W. van der Vaart(1998). Asymptotic Statistics. Cambridge University Press, UK.

L. Wang (2005). A Simple Nonparametric Test for Diagnosing Nonlinearity in Tobit Median Regression Model. Technical report. School of Statistics, University of Minnesota.

Z. Ying, S. H. Jung & L. J. Wei (1995). Survival analysis with median regression models. *Journal of the American Statistical Association,* 90 178-184.

S. Yang (1999). Censored median regression using weighted empirical survival and hazard functions. *Journal of the American Statistical Association,* 94 137-145.

S. G. Young & A. W. Bowman (1995). Non-parametric analysis of covariance. *Biometrics*, 51, 920-931.

J. X. Zheng (1998). A consistent nonparametric test of parametric regression models under conditional quantile restrictions. *Econometric Theory* 14, 123–138.

Received ???
Accepted ???

Lan WANG: `lan@stat.umn.edu`
School of Statistics, University of Minnesota
Minneapolis, MN 55455, USA